CONTENTS

---

3 Aggregate Loss Models

3.1 Introduction

Aggregate losses are the total losses paid by an insurer for a defined portfolio of insureds in one period, say a year. They are the sum of the individual losses for the year.

Definition 1.
The **collective risk model(Compound Distribution**, $S = X_1 + X_2 + \ldots + X_N$, is the aggregate loss variable, with N as the random variable representing the number of losses in the year. $X'_j s$ are independent and identically (i.i.d.) random variables. The independence assumptions are

(i) Conditional on $N = n$, the random variables $X_1 + X_2 + \ldots + X_n$ are i.i.d. random variables.

(ii) Conditional on $N = n$, the common distribution of the random variables $X_1 + X_2 + \ldots + X_n$ does not depend on n.

(iii) The distribution of N does not depend in any way on the values of $X_1 + X_2 + \ldots$.

3.2 The Recursive Method

Suppose that the severity distribution $f_X(x)$ is define on $0, 1, 2, \ldots, m$ representing multiples of some convenient monetary unit. The number m represents the largest possible payment and could be infinite. Further suppose that the frequency distribution, p_k, is a member of the $(a, b, 1)$ class and therefore satisfies

$$p_k = \left(a + \frac{b}{k}\right) p_{k-1}, k = 2, 3, 4, \ldots.$$

Then, the following result holds.

Theorem 1. For the $(a, b, 1)$ class,

$$f_S(k) = g_k = \frac{[p_1^M - (a+b)p_0^M]f_k + \sum_{j=1}^{k}\left(a + \frac{bj}{k}\right) f_j g_{k-j}}{1 - a f_0}$$

Corollary:

For the $(a, b, 0)$ class,

$$f_S(k) = g_k = \frac{\sum_{j=1}^{k}\left(a + \frac{bj}{k}\right) f_j g_{k-j}}{1 - a f_0}$$

The starting value of the recursive is

$$f_S(0) = P(S = 0)$$
$$= P_N[P_X(0)]$$
$$= P_N[f_0].$$

Example 1.

The number of claims in a period has a Negative Binomial distribution with parameters $r = 4$ and $\beta = 4$. The amount of each claim X follows $P(X = x) = 0.25$, $x = 0, 1, 2, 3$. The number of claims and claim amounts are independent. S is the aggregate claim amount in the period. Using recursive formula, calculate $F_S(3)$.

Example 2.

The number of claims on an insurance coverage follows a binomial distribution with parameters $m = 3$, $q = 0.2$. The size of each claim has the following distribution:

x	0	1	2
$P(X = x)$	0.50	0.35	0.15

Calculate the probability of aggregate claims of 3 or more. 0.0147

Example 3.

Taxis arrive at an airport in a Poisson process at a rate of 4 per minute. Each taxi picks up 1 to 4 passengers, with the following probabilities:

Number of Passengers	0	1	2	3
Probability, $P(X = x)$	0.7	0.2	0.05	0.05

Calculate the probability that in one minute 4 or more passengers leave the airport by taxi. 0.1671

Example 4.

For an insurance coverage, the number of claims has a zero-modified negative binomial distribution with parameters $r = 4$, $\beta = 1$ and $p_0^M = 0.5$. Claim size is distributed as follows:

Claim Size	1	2	3
Probability	0.50	0.40	0.10

Calculate $F_S(3)$. 0.63125

3.3 The Compound Model for Aggregate Claims: Convolution Method

The random sum, $S = X_1 + X_2 + \ldots + X_N$, has distribution function

$$F_S(x) = P(S \leq x)$$
$$= \sum_{n=0}^{\infty} p_n P(S \leq x | N = n)$$
$$= \sum_{n=0}^{\infty} p_n F_x^{*n}(x)$$

S had pdf

$$g_n = f_S(x) = \sum_{n=0}^{\infty} p_n f_X^{*n}(x),$$

for $x > 0$ if X is continuous

$$g_n = f_S(x) = \sum_{n=0}^{\infty} P(S = x | N = n),$$

for $x = 0, 1, \ldots$

Example 5.
Let $p_n = P[N = j - 1] = j/10$ for $j = 1, 2, 3, 4$, and let $f_X(1) = P(X = 1) = 0.4$, $f_X(2) = P(X = 2) = f_X(2) = 0.6$. $f_S(x)$ and $F_S(x)$ can be calculated as follows. $\square$

Example 6 (T3Q1).

The number of claims in a period has a Binomial distribution with parameters $m = 5$ and $q = 0.59$. The amount of each claim X follows $P(X = x) = 0.25$, $x = 1, 2, 3, 4$. The number of claims and claim amounts are independent. S is the aggregate claim amount in the period. Calculate $F_S(4)$.

Example 7 (T3Q2).

The number of claims in a period has a geometric distribution with mean 6.8. The amount of each claim is distributed as follows

Claim Amounts, X	0	1	2	3	4
Probability	0.37	0.34	0.12	0.10	0.07

The number of claims and claim amounts are independent. S is the aggregate claim amount in the period. Calculate $F_S(3)$.

Example 8 (T3Q3).

- Customers arrive in a store at a Poisson rate of 0.38 per minute.

- The amount of profit the store makes on each customer is randomly distributed as follows:

Profit	0	1	2	3
Probability	0.50	0.17	0.17	0.16

Determine the probability of making 3 profit in 10 minutes.

Example 9 (T3Q4).

Claim counts and sizes on an insurance coverage are independent and have the following distribution:

Number of claims	Probability
0	0.57
1	0.24
2	0.19

Claim Size	Probability
200	0.47
400	0.27
600	0.16
900	0.10

Let S be the aggregate claims. Calculate $F_S(600)$.

Example 10 (T3Q5).

The number of claims on an insurance coverage follows a zero modified Poisson distribution with mean $\lambda = 5$ and $p_0^M = 0.25$. The size of each claim has the following distribution:

Claim Size, x	0	3	6	9
Probability, $P(X = x)$	0.5	0.2	0.06	0.24

Calculate the probability of aggregate claims of 9 or more.

Example 11 (T3Q6).

The number of claims has a Poisson distribution with mean $\lambda = 1.8$. The distribution of the amount of claims(in thousand) is

Amount of claims	1	2	3	4	5	6
Probability	0.18	0.31	0.22	0.09	0.07	0.13

The number of claims and the amount of claims are independent. Determine the expected total amount of claims given that at least 4 thousand have been claimed. .

Example 12 (T3Q7)**.**

For an insurance coverage, you are given:

- Claim frequency (N^M), before application of deductibles, follows a zero modified geometric distribution with parameters $\beta = 7$ and $P(N^M = 0) = 0.6$.

- Claim size (X^M), before application of deductibles, follows a zero modified Poisson distribution with parameters $\lambda = 3$ and $P(X^M = 0) = 0.69$.

- Claim frequency and claim size are independent.

- There is a deductible of 3 per loss.

Calculate the probability that the number of payments is greater than 5.

Example 13 (T3Q8)**.**

Losses follow a compound distribution with both frequency and severity having discrete distribution.
For frequency

$$P_N(z) = 0.31 + 0.69 \left[\frac{(1 + 0.82(z - 1))^4 - (1 - 0.82)^4}{1 - (1 - 0.82)^4} \right]$$

For Severity

$$P_X(z) = 0.54 + 0.19z + 0.12z^2 + 0.08z^3 + 0.07z^4$$

Calculate the probability that aggregate losses is exactly 3.

3.4　Approximating Distribution

Calculating the distribution function for the aggregate distribution is difficult. An alternative is to use an approximating distribution. The aggregate distribution can be approximated with a normal distribution by the Central Limit Theorem when the sample size is large.

If the severity is discrete, then the aggregate loss distribution is discrete, and a continuity correction is required. This means adding or subtracting $1/2$ from the bound. For example, if you are asked for the probability that S is greater than 100, then we would calculate the probability that the normal variable is greater that 100.5, i.e.

- $P(S > s) = P(S > s + 0.5)$

- $P(S \geq s) = P(S \geq s - 0.5)$

- $P(S < s) = P(S < s - 0.5)$

- $P(S \leq s) = P(S \leq s + 0.5)$

$$
\begin{aligned}
E(S) &= E[E(S|N = n)] \\
&= E(X_1 + X_2 + \ldots + X_n) \\
&= E[nE(X)] \\
&= E(N)E(X) \\
&= \mu_N \mu_X
\end{aligned}
$$

The moments of S can be obtained in terms of the moments of N and the $X_j s$.

$$
\begin{aligned}
V(S) &= E[V(S|N = n)] + V[E(S|N = n)] \\
&= E[nV(X)] + V[nE(X)] \\
&= E(N)V(X) + V(N)E^2(X) \\
&= \mu_N \sigma_X^2 + \sigma_N^2 \mu_X^2
\end{aligned}
$$

By central Limit Theorem, S can be approximated by a normal distribution.

As $n \to \infty$,
$S \to N(\mu = \mu_N \mu_X, \sigma^2 = \mu_N \sigma_X^2 + \sigma_N^2 \mu_X^2)$

Probability Generating Function of S is given by

$$P_S(z) = P_N[P_X(z)]$$

The pgf is typically used when S is discrete.

Moment Generating Function:

$$M_S(z) = E(e^{zS}) = P_N[M_X(z)]$$

Example 14 (T3Q9)**.**

For insurance coverage, you are given:

- The number of claims for each insured follows a Binomial distribution with parameters $m = 5$ and q.

- q varies by insured according to beta distribution with parameters $a = 18$ and $b = 4$

- Claim size, before application to claims limits, follows a gamma distribution with parameters $\alpha = 4$, $\theta = 890$.

- Coverage is subject to claim limit of 1,970.

- Number of claims and claim sizes are independent.

Calculate the probability that aggregate losses will be greater than 1,532, using the normal approximation.

Example 15.

For a group insurance policy, the number of claims from a group has a binomial distribution with mean 100 and variance 20. The size of each claim has the following distribution:

Claim Size	probability
1	0.50
2	0.35
3	0.10
4	0.05

Using approximating normal distribution, calculate the probability that aggregate claims from this group will be greater than 180. $\boxed{0.1762}$

Example 16.

For aggregate losses $S = X_1 + X_2 + \ldots + X_N$, you are given:

(i) N has a Poisson distribution with mean 500.

(ii) $X_1, X_2, \ldots$ have mean 100 and variance 100.

(iii) $N, X_1, X_2, \ldots$ are mutually independent.

(iv) For a portfolio of insurance policies, the loss ratio is the ratio of aggregate losses to aggregate premiums collected.

(v) The premium collected is 1.1 times the expected aggregate losses.

Using the normal approximation to the compound Poisson distribution, calculate the probability that the loss ratio exceeds 0.95. $\boxed{0.1584}$

Example 17 (T3Q10)**.**

You are given:

- The number of claims for each insured follows a Poisson distribution with mean 6.

- Claim size, follows a Gamma distribution with parameters $\alpha = 5$, $\theta = 980$.

- Number of claims and claim sizes are independent.

Derive the formula for the skewness of the aggregate losses and then calculate the skewness.

Example 18 (T3Q11).

Losses follow a compound distribution with both frequency and severity having discrete distribution.
For frequency

$$P_N(z) = 0.21 + 0.79 \left[\frac{e^{2.60z} - 1}{e^{2.60} - 1} \right]$$

For Severity

$$P_X(z) = 0.55 + 0.21z + 0.11z^2 + 0.06z^3 + 0.07z^4$$

Calculate the Variance of the aggregate losses.

Example 19 (T3Q12).

For a certain insurance, individual losses in 2020 were Pareto distributed with parameters $\alpha = 4$ and $\theta = 1400$. A deductible of 140.0 is applied to each loss. In 2021, individual losses have increased 5%. A deductible of 140.0 is still applied to each loss. Determine the standard deviation of amount paid per loss.

3.5 Analytic Results

For certain combination of choices, simple analytic results are available, thus reducing the computational problems considerably.

3.5.1 Gamma Severity

If $X_j \sim gamma(\alpha, \theta)$, then

$$F_S(x) = \sum_{n=0}^{\infty} p_n F^{*n}(x)$$

Here, $F^{*n}(x) = P\left[\sum_{j=1}^{n} X_j \leq x \right]$

where $\sum_{j=1}^{n} X_j = S_{\alpha n} \sim gamma(\alpha^* = \alpha n, \theta)$

So,

$$F^{*n} = 1 - \sum_{j=0}^{\alpha n - 1} e^{-x/\theta} \frac{(x/\theta)^j}{j!}.$$

Substituting, we get

$$F_S(x) = \sum_{n=0}^{\infty} p_n \sum_{j=0}^{\alpha n - 1} \left(1 - \frac{(x/\theta)^j e^{-x/\theta}}{j!} \right)$$

$$= 1 - \sum_{n=1}^{\infty} p_n \sum_{j=0}^{\alpha n - 1} \left(\frac{(x/\theta)^j e^{-x/\theta}}{j!} \right)$$

$$= 1 - \sum_{n=1}^{\infty} p_n \sum_{j=1}^{\alpha n - 1} a_j$$

where $a_j = \frac{(x/\theta)^j e^{-x/\theta}}{j!}$ and $p_n = P[N = n]$

This is still an infinite sum, but if the frequency distribution is bounded, such as binomial, it is a finite sum.

$$F_S(x) = 1 - \sum_{n=1}^{m} \binom{m}{n} q^n (1-q)^{m-n}$$

$$\times \sum_{j=0}^{\alpha(n)-1} \left(\frac{(x/\theta)^j e^{-x/\theta}}{j!} \right)$$

Example 20.

Claim counts have a binomial distribution with $m = 2$, $q = 0.2$. Claim sizes are exponential with mean 1000. Calculate the probability that aggregate claims are less than their mean. 0.74796

Example 21.

Claim sizes follow a gamma distribution with parameters $\alpha = 2$ and $\theta = 100$. Claim counts are independent of claim sizes, and have the following distribution:

n	0	1	2
p_n	0.6	0.3	0.1

Calculate the probability that aggregate claims are less than 120. 0.7046

3.5.2   Compound Negative Binomial Exponential

$N \sim NB(r, \beta)$; $X \sim exp(\theta)$

$$M_S(z) = P_N[M_X(z)]$$

$$= P_N[(1 - \theta z)^{-1}]$$

$$= \{1 - \beta[(1 - \theta z)^{-1} - 1]\}^{-r}$$

$$= \left(1 + \tfrac{\beta}{1+\beta}\{[1 - \theta(1 + \beta)z]^{-1} - 1\}\right)^r$$

$$= P_N^*[M_X^*(z)]$$

where
$P_N^*(z) = [1 + \tfrac{\beta}{1+\beta}(z - 1)]^r$ and $M_X^*(z) = (1 - \theta(1 + \beta)z)^{-1}$

$$\implies N^* \sim Binomial\left[r, \tfrac{\beta}{1+\beta}\right] \text{ and } X^* \sim exp[\theta(1 + \beta)]$$

Thus, $F_S(x) = \displaystyle\sum_{j=0}^{\infty} p_n F^{*n}(x)$

where
$p_n = \binom{r}{n}\left(\tfrac{\beta}{1+\beta}\right)^n \left(\tfrac{1}{1+\beta}\right)^{n-r}$ and

$$F^{*n}(x) = P\left[\sum_{i=1}^{n} X_i^* \leq x\right]$$

where $\displaystyle\sum_{i=1}^{n} X_i^* \sim gamma(\alpha = n, \theta^* = \theta(1 + \beta))$

Thus,

$$F_S(x) = 1 - \sum_{n=1}^{r} \binom{r}{n}\left(\frac{\beta}{1+\beta}\right)^n \left(\frac{1}{1+\beta}\right)^{r-n}$$

$$\times \sum_{j=0}^{n-1} \frac{[x\theta^{-1}(1+\beta)^{-1}]^j e^{-x/[\theta(1+\beta)]}}{j!}$$

$$= 1 - \sum_{n=1}^{r} p_n \sum_{j=0}^{n-1} a_j$$

where $p_n = \binom{r}{n}\left(\tfrac{\beta}{1+\beta}\right)^n \left(\tfrac{1}{1+\beta}\right)^{r-n}$;

$a_j = \frac{[x\theta^{-1}(1+\beta)^{-1}]^j e^{-x/[\theta(1+\beta)]}}{j!}$

If $r = 1$, S has a compound geometric distribution, then

$$F_S(x) = 1 - \tfrac{\beta}{1+\beta}exp\left[-\tfrac{x}{\theta(1+\beta)}\right]$$

i.e. if $r = 1$, S is a two-point mixture of a degenerate distribution at 0 with weight $\tfrac{1}{1+\beta}$ and an exponential distribution with mean $\theta(1 + \beta)$, weight $\tfrac{\beta}{1+\beta}$.

Example 22 (T3Q13)**.**

Let the frequency distribution be negative binomial with $r = 4$ and $\beta = 4$. Let the severity distribution has the exponential distribution with mean 32. Determine $F_S(40)$

Example 23 (T3Q14)**.**

Claim sizes follow an exponential distribution with $\theta = 4.50$. Claim counts are independent of claim sizes, and have the following distribution:

n	0	1	2	3
P_n	0.39	0.29	0.18	0.14

Calculate $F_S(6)$.

Example 24.

Let the frequency distribution be negative binomial with $r = 2$ and $\beta = 2$. Let the severity distribution has the exponential distribution with mean 25. Determine $F_S(20)$ and $E(S \wedge 200)$. $\boxed{0.2284, 86.8754}$

3.6   Discretization:Method of Rounding

The recursive and convolution methods for calculating the aggregate distribution require a discrete distribution. Usually the severity distribution is continuous. We will pick a span, the distance between the points that will have a positive probability in the discretized distribution.

$f_0 = P(X < h/2) = F_X(h/2 - 0)$

$$f_{jh} = P(jh - h/2 \leq X < jh + h/2)$$

$$= F_X(jh + h/2 - 0) - F_X(jh - h/2 - 0),$$

$$= F_x\left[\frac{(2j+1)h}{2} - 0\right] - F_x\left[\frac{(2j-1)h}{2} - 0\right]$$

$$j = 1, 2, \dots$$

Notes:

- h is the span of the distribution

- -0 indicates that the lower bound is included but the upper bound isn't.

Example 25.
X has an exponential distribution with mean 1. Calculate p_2 of the distribution discretized using the method of rounding with a span of 1. $\boxed{0.1410}$

Example 26.
Loss sizes follow a Pareto distribution with $\alpha = 2$, $\theta = 3$. The distribution will be discretized by the method of rounding with a span of 4. Calculate the resulting probabilities of 0, 4, 8, and 12; f_0, f_4, f_8, and f_{12}. $\boxed{0.64, 0.24889, 0.05786, 0.02211}$

Example 27 (T3Q15)**.**

A random variable has an exponential distribution with mean 100. It is to be discretized using the method of rounding with span 90. Determine the mean of the discretized distribution.

3.7 Stop-loss Premium

Definition 2.

Insurance on the aggregate losses, subject to a deductible, is called stop-loss insurance. The expected cost of this insurance is called the net stop-loss premium and can be computed as $E[(S - d)_+]$, where d is the deductible and the notation $(.)_+$ means to use the value in parentheses if it is positive but use zero otherwise.

For any aggregate distribution,

$$E[(S - d)_+] = \int_d^\infty [1 - F_S(x)dx$$
$$= \int_d^\infty (x - d)f_S(x)dx$$
if X is continuous

$$E[(S - d)_+] = \int_d^\infty [1 - F_S(x)dx$$
$$= \sum_{k=d}^\infty (x - d)g_k$$
if X is discrete

Note that
$E[(S - d)_+] = E[S] - E[S \wedge d]$
Since $E[S] = E[N]E[X]$, we only have to deal with $E[S \wedge d]$.

3.7.1 Using the definition of $E[S \wedge d]$

For a discrete distribution in which the only possible values are multiples of h, it becomes

$$E[S \wedge d] = \sum_{j=0}^{[d/h]-1} h_j g_{hj} + dP[(S \geq d)]$$

Example 28 (T3Q16).

Prescription drug losses, S, are modeled assuming the number of claims has a geometric distribution with mean 6.0, and the amount of each prescription is 88.0. Calculate $E[(S - 220)_+]$.

Example 29 (T3Q17).

A company provides insurance to a concert hall for losses due to power failure. You are given:

- The number of power failures in a year has a Binomial distribution with parameters $m = 8$ and $q = 0.71$.

- The distribution of loss amount due to a single power failure is:

Loss Amount	10	20	30	40
Probability	0.50	0.20	0.17	0.13

- There is an annual deductible of 27.

Calculate the expected amount of claims paid by the insurer in one year.

Example 30.
You are given:

- The number of claims follows a binomial distribution with $m = 3$, $q = 0.2$.

- Claim sizes follow the following distribution:

Claim size	Claim probability
0	0.2
1	0.5
2	0.2
3	0.1

- A reinsurance policy has an aggregate deductible of 6.

Determine the expected aggregate amount paid by the reinsurer. $\boxed{0.000336}$

Example 31 (T3Q18)**.**

Claim counts follow a Poisson distribution with mean 1. Claim sizes follow an exponential distribution with $\theta = 700$. This severity distribution is discretized using the method of rounding with span 50. Claim counts and claim sizes are independent. A stop-loss reinsurance contract has a deductible of 110. Calculate expected losses paid by the reinsurance contract.

Example 32 (T3Q19)**.**

A company provides insurance to a concert hall for losses due to power failure. You are given:

- The number of power failures in a year has a Binomial distribution with parameters $m = 2$ and $q = 0.35$.

- The distribution of loss amount due to a single power failure follows a gamma distribution $\alpha = 2$ and $\theta = 11$.

- There is an annual deductible of 27.

Calculate the expected amount of claims paid by the insurer in one year.

3.7.2 Using Linear Interpolation

For the case where d is non-integral, we can use the following more general result that applies to continuous as well as discrete aggregate distributions.

Let the deductible d satisfy $a < d < b$, where $P(a < S < b) = 0$. Then $E[(S-d)_+]$ is obtained by linear interpolation between $E[(S-a)_+]$ and $E[(S-b)_+]$, producing

Theorem 2.

Suppose $P(a < S < b) = 0$. Then for $a \leq d \leq b$,

$$E[(S-d)_+] = \frac{b-d}{b-a} \cdot E[(S-a)_+] + \frac{d-a}{b-a} \cdot E[(S-b)_+].$$

Proof:

Example 33 (T3Q20)**.**

A stop-loss reinsurance pays 70% of the excess of aggregate claims above 1,030, subject to maximum payment of 357. For aggregate claims, S, you are given:

- $E[(S - 1,030)_+] = 530$

- $E[(S - 2,060)_+] = 265$

- The probability of an aggregate claim amount between 1,030 and 2,060 is zero.

Determine the total amount of claims the reinsurer expects to pay.

Example 34.
One way to establish the market view of probabilities is to consider the market price placed on contingent events in an efficient market. For example, suppose we have the following values from a competitive market:

$$E[(S - 5,000)_+] = 2000$$

$$E[(S - 10,000)_+] = 1500$$

Suppose also that loss amounts between 5,000 and 10,000 are impossible. According to this information, what is the probability that total losses will exceed 5,000? $\boxed{0.1}$